HUMAN SEMAPHORIN GA-1 (SEMAGA-A), A GENE INVOLVED IN NEURONAL DEVELOPMENT AND REGENERATION MECHANISMS DURING ADOPTOSIS, AND ITS USE AS A POTENTIAL DRUG TARGET

INSA1

Specification

The present invention relates to human semaphorin 6A-1 (SEMA6A-1), a novel gene involved in neuronal development and regeneration mechanisms during apoptosis.

Actin binding and filament assembly controlling proteins are essential for cellular events that require a drastic remodelling of cytoskeletal elements during development and apoptosis. Proline-rich proteins of the Ena/VASP family play a crucial role in actin and filament dynamics and have only recently been shown to be clustered to cell surface receptors like Dlar, a tyrosine phosphatase essential for motor axon outgrowth (F.B.Gertler et al., 1996, Cell 87, 227-239; Z.Wills et al., 1999, Neuron 22, 301-312). In the last decade the semaphorins were identified as a protein family displaying secreted or transmembrane-based repulsive guidance cues critically involved in neuronal development (J.G.Culotti and A.L.Kolodkin, Curr.Op.Neurobiol., 6, 81-88).

Therefore, it was an object of the present invention to provide a novel human semaphorin variant.

The invention comprises a nucleic acid coding for human semaphorin 6A-1 comprising

(a) the nucleotide sequence shown in SEQ ID NO:1,

(b) a sequence corresponding to the nucleotide sequence shown in SEQ ID NO:1 within the degeneration of the genetic code, or

(c) a sequence which hybridizes with the sequences of (a) or/and (b) under stringent conditions.

Surprisingly, the transmembranous human semaphorin 6A-1 ((HSA) SEMA6A-1) is capable of a selective binding to members of the Ena/VASP protein family. (HSA)SEMA6A-1 contains a cytoplasmic stretch at its C-terminal end. This domain shares a striking homology to Zyxin, a protein known to bind Ena/VASP (T.Macalma et al., 1996, JBC 271, 31470-31478; S.Hu and L.F.Reichardt, Neuron 22, 419-422). Thus, the human semaphorin sequence was found to comprise a section which matches with other semaphorin sequences, e.g. murine semaphorin sequences as well as a novel domain at its C-terminal end which is capable of binding to elements attached to the cytoskeleton.

Therefore, the invention further comprises a nucleic acid coding for a binding domain of human semaphorin 6A-1 comprising: (a) the nucleotide sequence shown in SEQ ID NO:3,(b) a sequence corresponding to the nucleotide sequence shown in SEQ ID NO:3 within the degeneration of the genetic code, or (c) a sequence which hybridizes with the sequences of (a) or/and (b) under stringent conditions.

The term "hybridization under stringent conditions" according to the present invention is used as described by Sambrook et al. (Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press (1989), 1.101-1.104). Preferably, a stringent hybridization according to the present invention is given when after washing for an hour with 1 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C, and most preferably at 68°C, and more preferably for 1 hour with 0.2 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C, and most preferably at 68°C a positive hybridization signal is still observed. A nucleotide sequence which hybridizes under such washing conditions with the nucleotide sequence shown in SEQ ID NO:1 or with a nucleotide

sequence corresponding thereto within the degeneration of the genetic code is a nucleotide sequence according to the invention.

The nucleic acid according to the invention preferably is in operative association with an expression control sequence that is active in eukaroytic cells, preferably in mammal cells.

The nucleotide sequence according to the invention preferably is a DNA. However, it may also be an RNA or a nucleic acid analog, such as a peptidic nucleic acid.

The nucleic acid according to the invention preferably comprises a sequence having a homology of greater than 80%, preferably greater than 90%, and more preferably greater than 95% and, in particular, greater than 97% to the nucleotide sequence according to SEQ ID NO:1. The term homology as used herein can be defined by the equation $H(\%) = [1-V/X] \cdot 100$, wherein H means homology, X is the total number of nucleobases of the nucleotide sequence according to SEQ ID NO:1 and V is the number of different nucleobases of a comparative sequence with regard to the nucleotide sequence according to SEQ ID NO:1.

The invention further comprises a polypeptide encoded by a nucleic acid according to the invention. Such a polypeptide is, in particular, capable of binding to members of the Ena/VASP protein family. The transmembranous SEMA6A-1 is capable of selectively binding to Evl but not Mena, both members of the Ena/VASP protein family.

The nucleic acids according to the invention can be obtained using known techniques, e.g. using short sections of the nucleotide sequence shown in SEQ ID NO:1 as hybridization probe or/and primer. They can, however, also be produced by chemical synthesis.

The invention further comprises a recombinant vector containing at least one copy of the nucleic acid according to the invention. This vector may be a prokaryotic or a eukaryotic vector which contains the nucleic acid according to the invention under the control of an expression signal (promoter, operator, enhancer etc.). Examples of prokaryotic vectors are chromosomal vectors such as bacteriophages and extra-chromosomal vectors such as plasmids, circulary plasmid vectors being particularly preferred. Prokaryotic vectors useful according to the present invention are, e.g., described in Sambrook et al., supra, chapter 1-4.

More preferably, the vector according to the invention is a eukaryotic vector, in particular a vector for mammal cells. Most preferred are vectors suitable for gene therapy, such as retrovirus, modified adenovirus or adeno-associated virus. Such vectors are known to the man skilled in the art of molecular biology and gene therapy and are also described in Sambrook et al., supra, chapter 16.

In addition to the polypeptide encoded by the nucleic acid of SEQ ID NO:1 or SEQ ID NO:3, the invention also relates to polypeptides differing therefrom by substitutions, deletions or/and insertions of single amino acids or short amino acid sections. The polypeptide is obtainable by expression of the nucleic acid sequence in a suitable expression system (cf. Sambrook et al., supra).

The polypeptide encoded by SEQ ID NO:1 is (HSA)SEMA6A-1, a new semaphorin variant containing a Zyxin-like domain that binds to the Ena/VASP-like protein (Evl). In particular, the semaphorins are a protein family displaying secreted or transmembrane-based repulsive guidance cues critically involved in neuronal development. The polypeptide encoded by SEQ ID NO:3 is a binding domain. This domain can bind selectively to Evl, a member of the Ena/VASP protein family. It may be particularly favorable to combine this binding domain with other proteins having known

functionality to give a fusion protein. This binding domain can be used advantageously, alone or as part of a fusion protein, as a means for screening and as a diagnostic and therapeutic target.

The invention further comprises a cell transformed with a nucleic acid or a vector according to the invention. The cell may be a eukaryotic or a prokaryotic cell, eukaryotic cells being preferred.

The present invention also comprises the use of the polypeptide or fragments thereof as immunogen for the production of antibodies. Standard protocols for obtaining antibodies may be used.

The present invention also comprises a pharmaceutical composition comprising a nucleic acid, modified nucleic acid, vector, cell, polypeptide or antibody as defined herein as active component.

The pharmaceutical composition may comprise pharmaceutically acceptable carriers, vehicles and/or additives and additional active components, if desired. The pharmaceutical composition can be used for diagnostic purposes or for the production of therapeutic agents. Particularly preferred is the use as a therapeutic agent for the modulation of the immune system.

Since the human semaphorin 6A-1 gene is involved in neuronal development and regeneration mechanisms during apoptosis, this gene can be used to design drug target structures. Members of the semaphorin gene family act as guidance signals and regulatory molecules during neuronal development. Besides its role in development, semaphorin has essential functions in the immune system. Semaphorin can also be linked to potential cancer, drug resistance and disease genes.

On the basis of a phylogenetic approach, the semaphorin gene family is currently distinguished into eight classes containing invertebrate (classes 1,

2) and vertebrate proteins (classes 3-7). Consistent with this nomenclature, the newly identified semaphorin is grouped into class 6 as human semaphorin 6A-1.

RNA expression studies have revealed SEMA6A-1 expression in areas consistent with a role of SEMA6A-1 as a guidance and regulatory signal during development and regeneration. Specialized domains in the cytoplasmic tail of the SEMA6A-1 gene product containing cytoskeletal binding elements show that SEMA6A-1 is also involved in differentiation, cytoskeletal stabilization and plasticity.

Finally, the invention is also directed to the use of the herein described pharmaceutical compositions for effecting differentiation, cytoskeletal stabilization and/or plasticity.

The invention is further described by the appended figures and examples, wherein

Figure 1 shows SEQ ID NO:1, the coding nucleotide sequence of the human semaphorin 6A-1 gene.

Figure 2 shows the nucleotide sequence of the human semaphorin 6A-1 gene as well as the derived amino acid sequence thereof;

Figure 3 shows the tissue distribution of (HSA)SEMA6A-1 revealed by Northern blot hybridizations of human embryo brain, lung, liver, kidney and human adult heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas tissue, respectively;

Figure 4 shows the (MMU)Sema6A-1 distribution in mouse adult and embryonic tissues revealed by in-situ hybridizations;

Figure 5 shows expression, protein size and dimerization of (HSA)SEMA6A-1;

Figure 6 shows a sequence alignment between SEMA6A-1 and Zyxin, wherein Figure 6a shows SEQ ID NO:3, the coding nucleotide sequence to a binding domain and Figure 6b shows the sequence of Zyxin;

Figure 7 shows immunoprecipitation of (HSA)SEMA6A-1 with α-Evl and α-Mena antibodies. A (α-Evl): Vector only (lane 1), pFlagSEMA6A-1 (lane 2), HT22 supplemented with purified SEMA6A-1 protein (lane 3), pFlagSEMA6A-1 precipitation using only protein A beads (lane 4), control detection of pFlagSEMA6A-1 transfected cells (lane 5), SEMA6A-1 purified control (lane 6), untransfected HT22 control (lane 7), Evl control in HT22 (lane 8); B (α-Mena): Vector only (lane 1), pFlagSEMA6A-1 (lane 2), HT22 supplemented with purified SEMA6A-1 protein (lane 3), control detection of pFlagSEMA6A-1 transfected cells (lane 4);

Figure 8 gives a graphical overview on the known Ena/VASP interacting proteins like Zyxin, Dlar and (HSA)SEMA6A-1.

Examples

Example 1

Cloning, genomic localization and tissue distribution of (HSA)SEMA6A-1

To identify and isolate repulsive guidance cues that might be involved in neuronal apoptosis a low stringency PCR-approach on cDNA from the human neuroblastoma cell line SK-N-MC was performed and a fragment of (HSA)SEMA6A-1 was amplified. This fragment was used to screen a human

1-ZAP Express cDNA library. Sequencing of 4 isolated clones revealed an ORF of 3093 bp referring to a protein of 1030 amino acids in total length with a predicted size of 135 kDa. (Fig.2: Nucleic acid sequence and deduced amino acid sequence).

Database searches identified 43 unordered sequences (Genbank Acc.-No. AC008524) and a mapped genomic survey sequence (Genbank Acc.-No. AB002453) of human chromosome 5 localizing the gene to 5q21-22. Gaps between the genomic sequences were closed by PCR on human genomic DNA and subsequent sequencing.

The (hsa)sema6A-1 gene covers 45 kb of genomic sequence and consists of 18 exons including 1 untranslated exon at the 3'-end (see Figure 2).

Example 2

Similarity and domain structure of (HSA)SEMA6A-1

Database searches revealed that SEMA6A-1 (1030aa) has a relatively high similarity to its murine ortholog Sema6A-1 (869aa) within the overlapping region consisting of 869aa. The existence of an additional cytoplasmic domain prompted us to name the new protein SEMA6A-1. This unique domain shares a 33% identity (49% similarity) to Zyxin, a proline-rich protein present at focal adhesion points and capable of binding to members of the Ena/VASP family. Binding of Zyxin to Ena/VASP occurs via a peptide stretch displaying the sequence DFPPPP (K.E.Prehoda et al., 1999, Cell 97, 471-480). (HSA)SEMA6A-1 contains two potential binding motifs (aa 858-962 (DNPPP) and aa 1010-1015 (DVPPKP) in its Zyxin homologous domain that are similar to the above-mentioned motif.

Example 3

Tissue distribution of (HSA)SEMA6A-1 revealed by Northern blot and in situ hybridization

Northern blot hybridizations of poly A^+ RNA of human adult and embryonic tissues detected two transcripts in the molecular range of 4.5 kb and 7 kb. Highest levels of detection were present in embryonic brain and kidney, moderate expression in lung and virtually no expression in liver. Compared to embryonic levels there was observed a clear reduction of expression of (HSA)SEMA6A-1 in adult tissues with the exception of placenta. In situ hybridizations in mouse embryo revealed a distinct expression throughout the whole embryo that is restricted to nervous system areas. These results indicate a general role of this protein in development and are shown in Figures 3 and 4: Figure 3 shows the human Northern blots. Figure 4 displays in situ hybridizations of embryonic (A, B, C, D) and adult (E, F, G) tissues. Notify the dominant expression in embryonic brain stem (A, B, D), optic precursors (A, C), spinal cord (B, D) and limb (B). High expression levels in adult regions are maintained in piriform cortex (E), cerebellar regions (F, G) and olfactory bulb (G).

Example 4

Expression of (HSA)SEMA6A-1 in mammalian cell lines

In order to show that Ena/VASP proteins might be potential intracellular binding partners for (HSA)SEMA6A-1 (see Figure 6, Alignment of (HSA)SEMA6A-1 and Zyxin) and that (HSA)SEMA6A-1 and Ena/VASP-like proteins might be interacting partners a XbaI/ScaI fragment of the SEMA6A-1 clone covering the full length protein sequence only lacking the signal sequence was subcloned into the pFLAG-CMV-1 vector. This vector allows rapid detection of the expressed fusion protein through the N-terminal Flag-Taq fused to the protein.

Immunoblotting of the tagged protein (Flag-SEMA6A-1) displayed a protein size of 125 kDa which closely corresponds to the predicted protein size. Expression in a human cell line (HEK293) and in a clonal mouse hippocampal cell line (HT22) followed by immunofluorescent analysis revealed that SEMA6A-1 is targeted to the cell surface and colocalizes with Evl and Mena, indicating a possible interaction between these proteins (see Figure 5, showing a graphical overview on the domain structure of (HSA)SEMA6A-1 and the subcloning strategy. In addition, Western blots displaying the protein size and its dimerization abilities are shown).

Example 5

Immunoprecipitation of (HSA)SEMA6A-1

Using antibodies specific for Mena and Evl Flag-SEMA6A-1 was immunoprecipitated from Triton X-100 extracts of transfected HEK239 and HT22 cells. The precipitate was separated by SDS-PAGE, and subsequent immunoblotting with the monoclonal anti-Flag antibody revealed that Flag-SEMA6A-1 co-immunoprecipitates with Evl but not Mena. To confirm this interaction Flag-SEMA6A-1 was purified from transfected HEK293 cells on an anti-Flag affinity column and the Triton X-100 extract of untransfected HT22 cells was supplemented with the purified protein, followed by immunoprecipiation of the protein complex using the α-Evl antibody. Immunoblotting again revealed that FlagSEMA6A-1 co-precipitates Evl. Figure 7 shows the immunoprecipitation experiments using the α-Evl- and α-Mena antibodies.